

```

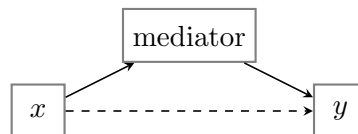
duration_cat  1758.0   3  4.8139 0.002960 **
sex            11.0   1  0.0902 0.764228
age            29.1   1  0.2392 0.625366
viq:age        41.6   1  0.3416 0.559586
Residuals     23372.7 192
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

```

6 Mediatie

6.1 Wat is mediatie?

- Mediatie = interventie:
 - De relatie tussen y en x wordt *gemedieerd* door een derde variabele
 - of nog: de mediator *interveniceert* in de relatie tussen y en x .



- Mediatie impliceert een conceptuele causale hypothese (*the mediational hypothesis*): de onafhankelijke variabele x beïnvloedt de mediator, en de mediator beïnvloedt de afhankelijke variabele y .
- Mediatie probeert te verklaren waarom x een invloed heeft op y .
- Mediatie-effecten zijn alomtegenwoordig in de gedragswetenschappen!
- In wat volgt beschouwen we zowel y als m (de mediator) als variabelen van minstens intervalniveau.

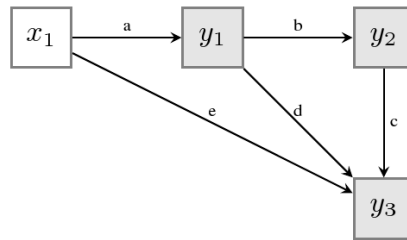
Een ‘klassieke’ paper waarbij het onderscheid tussen mediatie en moderatie aan bod komt, is:

Baron, R.M. & Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1185.

- Deze paper is al enorm veel geciteerd (zie Web of Science).
- De paper geeft een uitstekende beschrijving van het begrip *moderatie* en *mediatie*.
- Webpagina's:
 - <http://davidakenny.net/cm/moderation.htm>
 - <http://davidakenny.net/cm/mediate.htm>
- De auteurs beschrijven ook een ‘statistische procedure’ om na te gaan of er inderdaad sprake is van mediatie. Deze aanpak is bijzonder populair aangezien ze makkelijk uitvoerbaar is.
- In dit stuk nemen we de notatie van de auteurs over.

Paddiagrammen

- De schatting van een direct of rechtstreeks effect is een **padcoëfficiënt** (analoog met regressiecoëfficiënten)



- e representeert het direct effect van x_1 op y_3 .
- Het indirect effect van x_1 op y_3 :

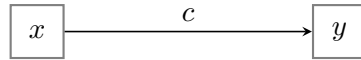
$$(a \times b \times c) + (a \times d)$$

- Het totaal effect van x_1 op y_3 : direct + indirect:

$$e + (a \times b \times c) + (a \times d)$$

Drie mogelijkheden m.b.t. mediatie

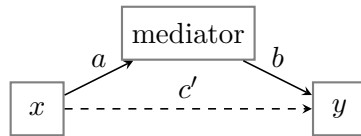
1. Geen mediatie: c = totaal effect van x op y :



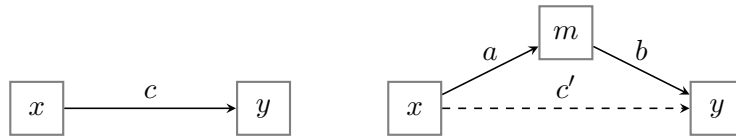
2. Volledige mediatie: $a \times b$ is het gemedieerd effect van x op y



3. Gedeeltelijke mediatie: *totaal* effect: $c' + a \times b$, *direct* effect: c'



6.2 De Baron & Kenny methode



1. Regresseer y op x : is er wel een verband tussen y en x ?

1. $H_0 : c = 0?$

2. Regresseer m op x : is er wel een verband tussen m en x ?

2. $H_0 : a = 0?$

3. Regresseer y op x en m : is er wel een effect van m op y na controle voor x ?

3. $H_0 : b = 0?$

4. Is er na controle van m nog wel een effect van x op y ?

4. $H_0 : c' = 0?$

- Er is sprake van volledige mediatie indien:
 1. $a \neq 0$, $b \neq 0$ en $c \neq 0$,
 2. $c' = 0$
- Er is sprake van gedeeltelijke mediatie indien:
 1. $a \neq 0$, $b \neq 0$ en $c \neq 0$,
 2. $c' < c$
- Bij deze mediatie-analyse maakt men de assumptie van lineaire relaties tussen de verschillende variabelen.
- Het verschil $(c - c')$ hanteert men vaak als een maat voor het *gemedieerd* effect.
- Voor lineaire modellen geldt: onder normale omstandigheden (geen missing values, zelfde covariaten in beide modellen)

$$(c - c') = a \times b.$$

- Stap 1 is eigenlijk overbodig: indien de mediator fungeert als een ‘suppressor’, dan zal er geen verband zijn tussen x en y , terwijl er wel sprake is van mediatie; een symptoom is dat het teken van $a \times b$ omgekeerd is aan c'

Een ‘suppressor’ of onderdrukkende variabele onderdrukt of verbergt de samenhang tussen 2 variabelen zodat deze geen verband met elkaar lijken te hebben.

Omwille van deze reden gebruiken we in deze cursus de Baron & Kenny methode als volgt: we toetsen het totale effect c in stap 1, maar ongeacht het resultaat gaan we over naar de volgende stappen, omdat er nog steeds sprake kan zijn van mediatie. We voeren dus altijd de 4 stappen uit en gaan op basis van de resultaten na of er (1) evidentie is voor mediatie (stappen 2 en 3) (2) indien wel, of er sprake is van gedeeltelijke of volledige mediatie of dat er sprake is van suppressie (stappen 1 en 4).

6.3 De Sobel test

Enkele problemen met de Baron & Kenny methode:

1. Conceptueel: het is een indirecte wijze om het mediatie effect na te gaan: we focussen op het verschil tussen c en c' , terwijl we eigenlijk geïnteresseerd zijn in a en b .
2. Statistisch:

- in vaak voorkomende situaties: weinig power (MacKinnon et al., 2002)
- meerdere toetsen na elkaar, dus een inflatie van Type I fouten
- mogelijkheid tot inconsistente resultaten over de verschillende regressies heen:
 - (a) a en b kunnen beide significant zijn, maar c' is niet kleiner dan c .
 - (b) c' is veel kleiner dan c , maar a en b zijn niet significant
- We bekomen enkel evidentie voor mediatie maar geen zekerheid; er zijn mogelijk andere modellen mogelijk! (bvb. m en y wisselen van plaats)
- Meer directe test: is het zo dat x een effect heeft op m ($= a$) en m een effect heeft op y ($= b$); met andere woorden, is $a \times b$ verschillend van nul?

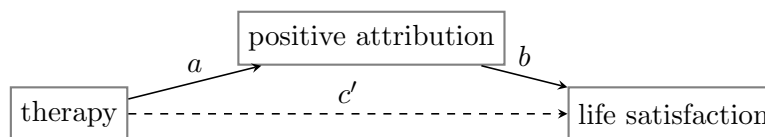
$$H_0 : a \times b = 0$$

Deze test is de Sobel test.

- Probleem: de steekproevenverdeling van $(a \times b)$ is doorgaans niet normaal verdeeld (vooral bij kleinere steekproefgroottes).
- Via bootstrap (subsamenen uit een dataset) is het mogelijk om een ‘empirische’ steekproevenverdeling voor $(a \times b)$ te berekenen.
- MacKinnon et al. (2004) hebben een exacte verdeling afgeleid voor de steekproevenverdeling van $(a \times b)$.

6.4 Voorbeeld

We beschouwen de fictieve data uit de paper van Preacher & Hayes (2004) waarbij men geïnteresseerd is in het effect van een nieuwe cognitieve therapie (**therapy**, nominaal) op de levenstevredenheid (**satisfaction**, van intervalniveau) na pensionering (zie sectie 1.2.4). De onderzoeksvraag is of het effect van de cognitieve gedragstherapie gemedieerd wordt door de positiviteit van de attributies (**attribution**, van intervalniveau).



We voeren de 4 stappen van de Baron & Kenny methode uit om na te gaan of er al dan niet sprake is van mediatie.

- Stap 1

```
> satis_fitxy<-lm(satis~therapy,data=satisfaction)
> summary(satis_fitxy)
```

Call:

```
lm(formula = satis ~ therapy, data = satisfaction)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.5669	-0.7319	0.3171	0.5121	1.3131

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.3271	0.2233	-1.465	0.1541
therapy	0.7640	0.3058	2.498	0.0186 *

Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Residual standard error: 0.8356 on 28 degrees of freedom

Multiple R-squared: 0.1823, Adjusted R-squared: 0.1531

F-statistic: 6.242 on 1 and 28 DF, p-value: 0.01862

De p -waarde van de toets voor $H_0 : c = 0$ (totaal effect van therapie op satisfactie) is gelijk aan 0.019; we kunnen H_0 bijgevolg verwerpen op het 5% significantieniveau.

- Stap 2

```
> satis_fitxm<-lm(attrib~therapy,data=satisfaction)
> summary(satis_fitxm)
```

Call:

```
lm(formula = attrib ~ therapy, data = satisfaction)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.4864	-0.5939	-0.0650	0.2611	1.8850

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
--	----------	------------	---------	----------

```
(Intercept)  -0.3536      0.2184  -1.619   0.1166
therapy       0.8186      0.2990   2.738   0.0106 *
```

```
---
```

```
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1
```

```
Residual standard error: 0.8171 on 28 degrees of freedom
```

```
Multiple R-squared:  0.2111, Adjusted R-squared:  0.183
```

```
F-statistic: 7.494 on 1 and 28 DF,  p-value: 0.01064
```

De p -waarde voor de toets $H_0 : a = 0$ (effect van therapie op attributie) is gelijk aan 0.011; we kunnen H_0 bijgevolg verwerpen op het 5% significantieniveau.

- Stap 3 + 4

```
> satis_fitxmy<-lm(satis~therapy+attrib,data=satisfaction)
> summary(satis_fitxmy)
```

```
Call:
```

```
lm(formula = satis ~ therapy + attrib, data = satisfaction)
```

```
Residuals:
```

```
Min      1Q   Median      3Q      Max
-1.36527 -0.60758  0.02416  0.54923  1.29091
```

```
Coefficients:
```

```
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.1843     0.2185  -0.844   0.406
therapy       0.4334     0.3221   1.346   0.190
attrib       0.4039     0.1808   2.234   0.034 *
```

```
---
```

```
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1
```

```
Residual standard error: 0.7818 on 27 degrees of freedom
```

```
Multiple R-squared:  0.3098, Adjusted R-squared:  0.2587
```

```
F-statistic: 6.06 on 2 and 27 DF,  p-value: 0.006697
```

De p -waarde voor de toets $H_0 : b = 0$ (effect van attributie op satisfactie, na controle voor therapie) is gelijk aan 0.034; we kunnen H_0 bijgevolg verwerpen op het 5% significantieniveau. Op basis van de tot nu toe uitgevoerde stappen kunnen we besluiten dat er volgens de Baron & Kenny methode aanwijzingen voor mediatie zijn.

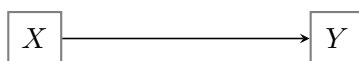
Verder zien we dat, na correctie voor attributie, het effect van therapie op satisfactie niet statistisch significant is op het 5% significantieniveau (p -waarde gelijk aan 0.19). We hebben dus evidentie voor volledige mediatie.

We kunnen afleiden dat $\hat{c} = 0.76$ (i.e. het geschatte totaal effect van therapie op satisfactie) en $\hat{c}' = 0.43$ (i.e. het geschatte effect van therapie op satisfactie na correctie voor attributie). Het geschatte gemedieerde effect is bijgevolg $0.76 - 0.43 = 0.33$.

We bekijken in een volgende stap enkel de resultaten voor de Sobel test. In R krijgen we voor de bootstrap methode het volgende 95% betrouwbaarheidsinterval voor $a \times b$: $[0.081, 0.803]$ (we gaan niet dieper in op de manier waarop deze resultaten bekomen worden). Er is geen sprake van mediatie indien $a \times b = 0$. Aangezien het betrouwbaarheidsinterval 0 niet omvat, kunnen we $H_0 : a \times b = 0$ verwerpen op het 5% significantieniveau en aannemen dat er sprake is van mediatie.

7 De ‘derde’ variabele

Veronderstel dat we geïnteresseerd zijn in het effect van een predictor X op een uitkomst Y .



Een andere predictor Z (de ‘derde’ variabele) kan X en/of Y of de relatie tussen X en Y op verschillende manieren beïnvloeden. We bekijken hier enkele mogelijkheden, merk op dat Z een set van predictoren kan voorstellen.

7.1 Confounding

Regressie wordt vaak gebruikt om causale verbanden te onderzoeken, maar causale verklaringen zijn niet zomaar gerechtvaardigd.

Regressiemodellen beschrijven associaties die niet noodzakelijk interpreteerbaar zijn als causale effecten. Causale besluitvorming is mogelijk onder specifieke assumpties die niet getest kunnen worden a.d.h.v. data maar die soms gegarandeerd worden door het design (denk aan experimentele versus observationele designs).

Voorbeeld